DOCUMENT RESUME

ED 136 242 CS 003 300

AUTHOR Baldwin, R. Scott
TITLE Psycholinguistic Strategies as a Factor in Estimating the Readability of Written Texts.
PUB DATE Apr 77
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, New York, April 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Connected Discourse; *Context Clues; Grade 3; *Miscue Analysis; Oral Reading; Primary Education; Psycholinguistics; *Readability; Reading Comprehension; *Reading Research; Silent Reading; *Syntax

ABSTRACT

Fifty-six third graders were randomly assigned to two treatment groups, in a study of the relationship between clause structure and the readability of written texts. The treatment groups read sets of passages which were identical except for certain word-order modifications. The dependent variables were silent- and oral-reading comprehension, rate of oral reading, and number of intonation miscues generated at clause boundaries. Treatment groups differed at the .01 level in silent-reading comprehension and number of intonation miscues; no significant differences were obtained for rate or comprehension during oral reading. Specific clause-analysis strategies derived from speech perception theory were proposed, to account for the experimental effects. It was concluded that clause-analysis strategies should be a factor in explaining the syntactic complexity of written materials. (Author/AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED136242

Psycholinguistic Strategies as a Factor in Estimating the Readability of Written Texts

R. Scott Baldwin

The University of Tulsa

Running head: Psycholinguistic Strategies...Readability

CS 003 300

Abstract

The study investigated the relationship between clause strategies and the readability of written texts. Fifty-six third graders were randomly assigned to two treatment groups. The treatment groups read sets of passages which were identical except for certain word order modifications. The dependent variables were silent and oral reading comprehension, rate of oral reading, and number of intonation miscues generated at clause boundaries. At $p < .01$, treatment groups were found to differ in silent reading comprehension and number of intonation miscues. Statistically nonsignificant differences were obtained for rate and comprehension during oral reading. Specific clause analysis strategies derived from speech perception theory were proposed to account for the experimental effects. It was concluded that clause analysis strategies should be a factor in explaining the syntactic complexity of written materials.

Reading specialists and researchers have long known that variation in written syntax affects reading comprehension. Traditionally, readability assessments have utilized sentence length as a predictor of syntactic complexity, and more recently, attempts have been made to explain the syntactic complexity of written materials in terms of derivational complexity, Coleman (1964), Fagan (1971), and Evans (1972), and the extent to which written sentences match the oral language patterns of the reader, Ruddell (1965), Tatham (1970), and Reid (1972). The general thesis of this research is that an additional factor, the actual processing strategies of the readers themselves, is requisite to a more complete explanation of the effects of syntax upon reading comprehension.

According to Fodor, Bever, and Garrett (1974), speech perception involves the gathering of information regarding the nature of sentoids, deep structure clauses, which are usually coterminous with surface structure clauses. One heuristic means of gathering such information is the canonical-sentoid strategy, which refers to the hypothesis that the hearer assumes initially that a noun phrase immediately followed by a verb phrase immediately followed by a noun phrase refer, respectively, to the subject, main verb, and object of a common deep structure clause.

If such heuristic strategies exist in speech perception, it is logical to question whether the same or similar strategies are part of the reading process. Specifically, the present investigation attempted to question the existence of one particular set

of syntactic strategies as well as to determine their collective relationship to the complexity of written materials.

Definition of Terms

"Canonical" refers to a fundamental schema which is the clearest or simplest possible, in this case a grammatical schema; and sentoid, within the context of this study, refers to any clause.

There were two clausal shapes or schemas which were included in this study. The first is a statement with the basic grammatical form NP^1-V-(NP^2), where NP^1 refers to the clausal subject, V refers to the main verb of the clause, and (NP^2) refers to an object which is optional, depending upon whether or not the main verb is transitive. The second clausal shape is a question with the basic grammatical form V-NP-X, where V refers to a copula or auxiliary verb, NP refers to the subject of the clause, and X refers to anything that follows the subject of the clause.

Consider the following matched pairs of sentences:

1. (a) John$_\downarrow$ come here.
   (b) Come here$_\downarrow$ John.

2. (a) Tim$_\downarrow$ will you help me?
   (b) Will you help me$_\downarrow$ Tim?

3. (a) Ralph said$_\downarrow$ "He is a dingbat."
   (b) "He is a dingbat$_\downarrow$" said Ralph.

4. (a) Why can't we eat, John?

   (b) John, why can't we eat?

5. (a) If you can, skip along.

   (b) Skip along, if you can.

6. (a) Get on. The bike is fun.

   (b) Get on the bike. It is fun.

7. (a) "Mary can go?" asked Pete.

   (b) "Can Mary go?" asked Pete.

8. (a) "Are you funny!" said Aphrodite.

   (b) "You are funny!" said Aphrodite.

Each of the above pairs are equivalent semantically and in terms of sentence structure complexity, i.e., they mean the same thing, and no presently available method of measuring sentence structure complexity has the capacity to distinguish between the respective (a)'s and (b)'s. However, this study attempted to substantiate the claim that for some readers such sentence pairs differ internally from the standpoint of sentence structure complexity. In all eight cases, (a) is radically more complex than (b). The complexity in each instance is indicated by the fact that in (a) the underlined punctuation mark is critical to the correct identification of the sentence's grammatical structure. The same is not true of (b), where the punctuation mark is never critical and in some cases is completely redundant. For instance,

in example 1(a) the comma cues the reader that "John" is a noun of direct address; but if the comma is deleted or simply not observed, the sentence takes on the appearance of a simple statement with "John" functioning as the sentence subject. This potential for ambiguity does not exist in (b) if the comma is deleted. A similar analysis can be made for the other seven pairs.

One of the interesting facets of this phenomenon is the fact that a purely linguistic approach is insufficient for explaining the differences between the above sentence pairs. In speech, any matched (a) and (b) are equally appropriate and are probably equally easy to produce or perceive, i.e., the transformational history of (a) type sentences is generally no more complex than that of (b) type sentences. However, in the reading situation, this equivalence of complexity is not maintained because (b) type structures conform to the expectations of novice readers; (a) type structures do not. For this reason, type (a) clause structures will be labeled "noncanonical" and type (b) structures will be termed "canonical".

The theory underlying this investigation is that readers, at least during the acquisition stage, actively employ abstract clause configurations as guides in identifying and predicting the grammatical structures of written sentences. For example, the reader will tend to identify the initial NP of a clause as the subject of that clause. The identification of the clausal subject causes the reader to anticipate a following verb and object. Studies investigating grammatical structure in children's

language and in materials written for primary level readers, e.g., O'Donnell, Griffin, and Norris (1967), have revealed that the simple subject-verb-object construction is the most frequently occurring of grammatical patterns. Consequently, the prediction that children search for this pattern is neither profound nor surprising. What is more interesting is the prediction that clause strategies are so strong in some readers that they will seek to confirm their own expectations even in the face of overt visual cues to the contrary, i.e., punctuation.

Method

Fifty-six third grade students from a rural elementary school in Southeastern Ohio were randomly assigned to two treatment groups. The stimulus materials in the experiment consisted of two passages written by the investigator. Each passage was approximately 300 words in length, was written on a second grade level, and had two alternate forms which were randomly assigned to treatment groups. One form of each passage contained noncanonical, type (a), clause configurations. The other form contained canonical, type (b), clause structures which were identical to the corresponding noncanonical structures in vocabulary, length, and number of transformations required to derive them. Each treatment group read one passage orally and one silently. The oral reading was timed and also tape recorded so that intonation miscues at clause boundaries could be

evaluated. Following the reading of each passage, a twelve item multiple choice comprehension test was administered.

Results

A multivariate analysis of variance was performed in order to assess differences between treatment groups on the following dependent variables: Comprehension during silent reading, comprehension during oral reading, speed of oral reading, and intonation miscues at selected clause boundaries.

At the $p < .01$ level of significance, treatment groups were found to differ on silent reading comprehension and on the number of intonation miscues generated at clause boundaries during oral reading. In both cases, the performance of the group reading passages containing type (a) sentences was poorer. Nonsignificant differences were obtained on oral reading comprehension and speed during oral reading.

Multivariate Test for Equality of Mean Vectors, F 4, 51 = 21.93, $p < .0001$

Univariate ANOVAS

Dependent Variable	Source of Variation	df	MS	F	P
Oral Reading Comprehension	Canonical/Non-canonical	1	2.16	.92	.3429
	Within	54	2.36		
Silent Reading Comprehension	Canonical/Non-canonical	1	19.45	7.52	.0083
	Within	54	2.59		
Intonation Miscues	Canonical/Non-canonical	1	440.16	62.48	.0001
	Within	54	7.04		
Rate of Reading	Canonical/Non-canonical	1	721.45	.21	.6448
	Within	54	3356.66		

Multivariate and Univariate Analyses of Variance for the Two Cell Design Comparing Reading Comprehension Under Canonical and Non-canonical Clause Conditions

Discussion

The results of the study warrant several conclusions. First, the difference in reading comprehension between treatment groups under the silent condition suggests that the type (a) passage was, in some sense, more complex than its type (b) counterpart. And I infer from this that the type (a) passage had a higher level of readability. Since the passages were 100% identical

in terms of vocabulary content and sentence length, any readability formula or graph that I am aware of will generate identical indexes of difficulty for the two passages; and, to the best of my knowledge, matched clauses and sentences within the passages are generated by the same number and types of transformations. In addition, I am unaware of any evidence which would support the notion that type (a) clause structures constitute a bizarre or infrequent phenomenon in the speech of third grade children. Apparently, another factor is responsible for the treatment effects. I believe that this factor is the syntactic processing strategies of the readers themselves. Consider the responses to Item 7 of the silent reading comprehension test.

Question #7: When they got to the zoo, where were the animals?

Contexts to which the question referred:

(1a) In the type (a) passage, "The sun was shining and the animals were all outside. The zoo had many visitors."

and

(1b) In the type (b) passage, "The animals were all outside, and the sun was shining. The zoo had many visitors."

Multiple-choice answers:

(a) in their cages (b) out in the sunshine

(c) outside the zoo (d) under a tree

On this particular test item, nine subjects reading the non-canonical type (a) passage but only two subjects reading the

canonical type (b) passage responded by answering (c), "outside the zoo."

The second sentence in (1a) is non-canonical because it may falsely confirm the reader's expectation for an object in the initial sentence. On the other hand, the second sentence in (1b) is canonical because it does not falsely confirm any syntactic expectations generated in the initial sentence. If subjects were actually employing the hypothesized clause strategies, then some of the subjects reading the type (a) passage probably failed to "observe" the period between outside and the zoo. Accordingly, I am hypothesizing that some subjects converted (1a) into (1c).

(1c) The sun was shining, and the animals were all outside the zoo. The zoo had many visitors.

Similar patterns of response were evident in subjects' answers to other test items, thus supporting the clause strategy hypothesis. Grammatical complexity in reading appears, in part, to be a function of the match between predicted and actual clause structures in the text.

A second conclusion of the study is that syntactic processing is different in oral and silent reading. Non-canonical structures seemed to have no negative affect upon comprehension during oral reading in spite of the fact that subjects reading the type (a) passage made massive numbers of intonation errors at clause boundaries.

The fact that third graders failed to attend to punctuation and their corresponding clause boundaries in the environment of

type (a) structures should be brought to the direct attention of reading diagnosticians and teachers. The beginning levels of basals tend to include large numbers of type (a) grammatical structures (Coady and Baldwin, 1975); and since type (a) structures can inhibit reading comprehension, their wide use in primary level reading materials is altogether inconsistent with one of the most basic of educational procedures: introduce that which is simple and then follow it with that which is complex. Type (a) structures can usually be converted with ease into type (b) structures, which are apparently more consistent with the syntactic processing strategies of young readers. Just as vocabulary and content should be introduced to the novice with care so that learning is positively facilitated, so too should type (b) clauses and sentences be introduced before their type (a) counterparts. In addition, where type (a) structures exist, teachers should be made aware of the fact that their punctuation is critical to correct syntactic processing.

The results of this experiment are not altogether unambiguous. For instance, I am suspicious of the contradictory results under the oral and silent conditions. I find it counterintuitive that syntactic processing should be so radically different during oral and silent reading. The fact that different texts were used in the oral and silent treatments could account for the discrepency in treatment effects.

In spite of the obvious limitations of the present study, I think there is one major implication for the study of readability.

Simple taxonomies of child speech patterns and the direct application of transformational grammar cannot fully explain the notion of syntactic complexity, at least not in reading. Sooner or later we will have to begin describing the precise strategies which readers use, because "complexity", grammatical or otherwise, is a function of what is in the reader's head and not of what is on the page in front of him.

References

Coady, J. M., and Baldwin, R. S. Syntax and primers. to appear in Reading Improvement.

Coleman, E. B. The comprehensibility of several grammatical transformations. Journal of Applied Psychology, 1964, 48, 186-190.

Evans,R.V. The effects of transformational simplification on the reading comprehension of selected high school students. Journal of Reading Behavior, 1972, 5, 273-281.

Fagan,W.T. Transformations and comprehension. The Reading Teacher, 1971, 25, 169-172.

Fodor, J. A., Bever, T. G., and Garrett, M. F. The psychology of language. New York: McGraw-Hill Book Company, 1974.

O'Donnell, R. C., Griffin, W. J., and Norris, R. C. Syntax of kindergarten and elementary school children: A transformational analysis. Champaign, Illinois: NCTE, 1967.

Reid, J. F. Children's comprehension of syntactic features found in some extension readers. In Jessie F. Reid (ed), Reading: problems and practices. London: Ward Lock, 1972, pp. 394-403.

Ruddell, R. B. Effect of the similarity of oral and written patterns of language structure on reading comprehension. Elementary English, 1965, 42, 403-410.

Tatham, S. M. Reading comprehension of materials written with select oral language patterns. Reading Research Quarterly, 1970, 5, 402-426.